CSE508 Information Retrieval - Winter 2024

Om Merholiya
Roll No: 2021404

1 Question 1: Image and Text Feature Extraction

1.1 1.1 Overview

Our approach for question 1 integrates both image and text data for feature extraction, leveraging pre-trained models and TF-IDF vectorization to create a rich feature set.

1.2 1.2 Methodologies

For images, we use ResNet50 to extract deep features, and for text, we employ TF-IDF vectorization to capture term significance across documents.

1.3 1.3 Assumptions

We assume that our dataset's images are accessible and preprocessable, and that TF-IDF effectively captures the textual data's essence.

1.4 1.4 Results

The implementation demonstrates successful feature extraction, setting a solid foundation for subsequent retrieval tasks.

2 Question 2: Similarity Calculation and Retrieval

2.1 2.1 Overview

Question 2 focuses on calculating cosine similarities for both image and text features to facilitate effective retrieval.

2.2 2.2 Methodologies

We calculate cosine similarities separately for the image and text features, then merge these to identify the most relevant items for a given query.

2.3   2.3 Assumptions

It's assumed that the extracted features are representative enough to calculate meaningful similarities.

2.4   2.4 Results

The cosine similarity scores allow for identifying closely related items to the query, proving the effectiveness of our feature representation.

3   Question 3: Combined Retrieval Using Composite Scores

3.1   3.1 Overview

This question introduces the concept of composite similarity scores, combining image and text similarities for improved retrieval.

3.2   3.2 Methodologies

A composite score is calculated by averaging the image and text similarity scores, used to rank and retrieve items.

3.3   3.3 Assumptions

The composite score is presumed to offer a balanced measure of similarity, equally weighting text and image features.

3.4   3.4 Results

The approach successfully retrieves relevant items, highlighting the value of integrating multiple data modalities.

4   Conclusion

This review summarizes our methodologies and findings across the assigned tasks, demonstrating the potential of multimodal information retrieval. Fu-

ture work may explore refining similarity measures and integrating user feedback for enhanced retrieval performance.

5 Question 4: Composite Similarity Scores and Ranking

5.1 Overview

Question 4 focused on computing composite similarity scores by averaging image and text similarity scores obtained in the previous steps, and then ranking the items based on these composite scores.

5.2 Methodologies

Composite similarity scores were calculated by averaging image and text similarity scores. The ranked indices were determined based on these composite scores, providing a combined measure of similarity for each item.

5.3 Assumptions

The primary assumption is that averaging image and text similarity scores provides a balanced measure of overall similarity. Additionally, it is assumed that the ranking based on composite scores accurately reflects the relevance of items to the input query.

5.4 Results

The calculated composite similarity scores and ranked indices facilitated the identification of the top items most similar to the input query. The ranking based on composite scores allowed for a comprehensive assessment of item relevance, considering both visual and textual content.

6 Question 5: Retrieval Results and Analysis

6.1 Overview

Question 5 involved analyzing the retrieval results obtained in previous steps, particularly focusing on the relationship between image and text similarity scores and their impact on composite similarity scores.

6.2 Methodologies

Retrieval results were analyzed by examining the relationship between image and text similarity scores and assessing their influence on composite similarity scores. Comparisons were made to determine if the highest-ranked images corresponded to the highest-ranked text and vice versa.

6.3 Assumptions

The primary assumption is that higher similarity scores indicate a stronger match between items. Additionally, it is assumed that a consistent relationship exists between image and text similarity scores.

6.4 Results

The analysis revealed insights into the relationship between image and text similarity scores and their combined effect on composite similarity scores. Comparisons between image and text rankings provided valuable insights into the effectiveness of the retrieval process and the relevance of items to the input query.

7 Question 6: User Input Evaluation

7.1 Overview

Question 6 focused on evaluating the retrieval system using user input, assessing its effectiveness in returning relevant items based on both image and text queries.

7.2 Methodologies

User input was provided in the form of an image URL and review text. The retrieval system processed the input and returned the most similar items based on both image and text features. Evaluation was performed based on the relevance and accuracy of the retrieved items.

7.3 Assumptions

The primary assumption is that user input accurately reflects their information needs, and the retrieval system effectively captures these needs. Additionally, it is assumed that the retrieved items provide relevant information corresponding to the input query.

7.4 Results

The evaluation of the retrieval system using user input demonstrated its effectiveness in returning relevant items based on both image and text queries. The retrieved items closely matched the input query, indicating the system's capability to accurately identify relevant content.